

Ishimura, K. and Nakada, H.
Osaka J. Math.
44 (2007), 667–683

SOME ERGODIC PROPERTIES OF THE NEGATIVE SLOPE ALGORITHM

KOSHIRO ISHIMURA and HITOSHI NAKADA*

(Received June 12, 2006, revised October 23, 2006)

Abstract

The notion of the negative slope algorithm was introduced by S. Ferenczi, C. Holton, and L. Zamboni as an induction process of three interval exchange transformations. Then S. Ferenczi and L.F.C. da Rocha gave the explicit form of its absolutely continuous invariant measure and showed that it is ergodic. In this paper we prove that the negative slope algorithm with the absolutely continuous invariant measure is weak Bernoulli. We also show that this measure is derived as a marginal distribution of an invariant measure for a 4-dimensional (natural) extension of the negative slope algorithm. We also calculate its entropy by Rohlin's formula.

1. Introduction

The negative slope algorithm (n.s.a.) was introduced by S. Ferenczi, C. Holton, and L. Zamboni [1] to discuss the structure of three-interval exchange transformations, see also [2] and [3]. It is a kind of multidimensional continued fractions algorithm and some arithmetic properties were discussed in [1]. Recently, S. Ferenczi and L.F.C. da Rocha [4] discussed its ergodic properties. Indeed, they showed the existence of an absolutely continuous invariant measure, which is ergodic. In this paper, we show that the n.s.a. satisfies conditions given by M. Yuri [7], which imply Rohlin's entropy formula and weak Bernoulli property with respect to the absolutely continuous invariant measure. We also derive the absolutely continuous invariant measure given in [4] from a 4-dimensional representation of the natural extension of the n.s.a. and compute the explicit value of entropy of the n.s.a. by Rohlin's entropy formula and give the exponent constant of the denominator of the n -th convergent of simultaneous approximations arising from the n.s.a. In §2, we give the definition of the n.s.a. and some basic notions related to the n.s.a. Then, in §3, we explain some sufficient conditions by [7] for multi-dimensional maps T to be weak Bernoulli. In §4, we prove Rohlin's formula and the weak Bernoulli property of the n.s.a. by showing a number of properties which implies that Yuri's condition holds for the n.s.a. Finally, in §5, we construct a 4-dimensional map, which is the natural extension of the

2000 Mathematics Subject Classification. 11K55, 37A05, 37A25.

*Partially supported by Grant-in Aid for Scientific research (No. 18340032) by Japan Society for the Promotion of Science.

n.s.a. and derive the absolutely continuous invariant measure for the n.s.a. as the marginal distribution. Then we calculate the entropy of the n.s.a. by Rohlin's formula.

2. Basic notions of the negative slope algorithm

First we define a map T on the unit square, which is called the negative slope algorithm. For $(x, y) \in \mathbb{X} = [0, 1]^2 \setminus \{(x, y) \mid x + y = 1\}$, we define

$$T(x, y) = \begin{cases} \left(\frac{y}{(x+y)-1} - \left\lfloor \frac{y}{(x+y)-1} \right\rfloor, \frac{x}{(x+y)-1} - \left\lfloor \frac{x}{(x+y)-1} \right\rfloor \right) & \text{if } x + y > 1 \\ \left(\frac{1-y}{1-(x+y)} - \left\lfloor \frac{1-y}{1-(x+y)} \right\rfloor, \frac{1-x}{1-(x+y)} - \left\lfloor \frac{1-x}{1-(x+y)} \right\rfloor \right) & \text{if } x + y < 1. \end{cases}$$

We put

$$n(x, y) = \begin{cases} \left\lfloor \frac{y}{(x+y)-1} \right\rfloor & \text{if } x + y > 1 \\ \left\lfloor \frac{1-y}{1-(x+y)} \right\rfloor & \text{if } x + y < 1, \end{cases}$$

$$m(x, y) = \begin{cases} \left\lfloor \frac{x}{(x+y)-1} \right\rfloor & \text{if } x + y > 1 \\ \left\lfloor \frac{1-x}{1-(x+y)} \right\rfloor & \text{if } x + y < 1, \end{cases}$$

and

$$\varepsilon(x, y) = \begin{cases} -1 & \text{if } x + y > 1 \\ +1 & \text{if } x + y < 1. \end{cases}$$

Then we put

$$\begin{cases} n_k(x, y) = n(T^{k-1}(x, y)) \\ m_k(x, y) = m(T^{k-1}(x, y)) \\ \varepsilon_k(x, y) = \varepsilon(T^{k-1}(x, y)) \end{cases}$$

for $k \geq 1$. Then we have a sequence

$$((\varepsilon_1(x, y), n_1(x, y), m_1(x, y)), (\varepsilon_2(x, y), n_2(x, y), m_2(x, y)), \dots)$$

for each $(x, y) \in \mathbb{X}$. We note that $n_k, m_k \geq 1$ for $k \geq 1$ and for any sequence $((\varepsilon_i, n_i, m_i), i \geq 1)$, there exists $(x, y) \in \mathbb{X}$ such that $(\varepsilon_i(x, y), n_i(x, y), m_i(x, y)) = (\varepsilon_i, n_i, m_i)$ unless there exists $k \geq 1$ such that $(\varepsilon_i, m_i) = (+1, 1)$ for $i \geq k$ or $(\varepsilon_i, n_i) = (+1, 1)$ for $i \geq k$. By [1], we see that

$$\{(\varepsilon_k(\cdot), n_k(\cdot), m_k(\cdot)), k \geq 1\}$$

separates points. Indeed, [1] showed that the sequence of digits separates $\{(x + y, x - y)\}$, which is equivalent to the above, i.e. if $(x, y) \neq (x', y') \in \mathbb{X}$, then there exists $k \geq 1$ such that

$$(1) \quad (\varepsilon_k(x, y), n_k(x, y), m_k(x, y)) \neq (\varepsilon_k(x', y'), n_k(x', y'), m_k(x', y')).$$

Now we introduce the projective representation of T . We put

$$A_{(+1,n,m)} = \begin{pmatrix} n & n-1 & 1-n \\ m-1 & m & 1-m \\ -1 & -1 & 1 \end{pmatrix}$$

and

$$A_{(-1,n,m)} = \begin{pmatrix} -n & -n+1 & n \\ -m+1 & -m & m \\ 1 & 1 & -1 \end{pmatrix}$$

for $m, n \geq 1$. Then we see

$$A_{(+1,n,m)}^{-1} = \begin{pmatrix} 1 & 0 & n-1 \\ 0 & 1 & m-1 \\ 1 & 1 & n+m-1 \end{pmatrix}$$

and

$$A_{(-1,n,m)}^{-1} = \begin{pmatrix} 0 & 1 & m \\ 1 & 0 & n \\ 1 & 1 & n+m-1 \end{pmatrix}.$$

We identify (x, y) to $\begin{pmatrix} \alpha x \\ \alpha y \\ \alpha \end{pmatrix}$ for $\alpha \neq 0$. Then $T(x, y)$ is identified to

$$A_{(\varepsilon_1(x,y), n_1(x,y), m_1(x,y))} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

and its local inverse is given by

$$A_{(\varepsilon_1(x,y), n_1(x,y), m_1(x,y))}^{-1}.$$

In this way, we get a representation of (x, y) by

$$A_{(\varepsilon_1, n_1, m_1)}^{-1} A_{(\varepsilon_2, n_2, m_2)}^{-1} A_{(\varepsilon_3, n_3, m_3)}^{-1} \cdots$$

and T is defined as a multiplication by $A_{(\varepsilon_1, n_1, m_1)}$ from the left and acts as a shift on the set of infinite sequence of matrices

$$\{A_{(\varepsilon_1, n_1, m_1)}^{-1} A_{(\varepsilon_2, n_2, m_2)}^{-1} A_{(\varepsilon_3, n_3, m_3)}^{-1} \cdots \mid \varepsilon_k = \pm 1, n_k, m_k \geq 1 \text{ for } k \geq 1\}.$$

For a given sequence $((\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k))$, we define a cylinder set of length k by

$$\begin{aligned} & \langle (\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k) \rangle \\ &= \{(x, y) \mid (\varepsilon_i(x, y), n_i(x, y), m_i(x, y)) = (\varepsilon_i, n_i, m_i), 1 \leq i \leq k\}. \end{aligned}$$

For $(x, y) \in \langle (\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k) \rangle$, $T^k(x, y)$ is expressed as

$$A_{(\varepsilon_k, n_k, m_k)} \cdots A_{(\varepsilon_1, n_1, m_1)} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

and its local inverse $\Psi_{\langle (\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k) \rangle}$ is expressed as

$$A_{(\varepsilon_1, n_1, m_1)}^{-1} \cdots A_{(\varepsilon_k, n_k, m_k)}^{-1}.$$

Since

$$\begin{aligned} & \left\{ \left(\frac{y}{(x+y)-1}, \frac{x}{(x+y)-1} \right) : (x, y) \in \mathbb{X}, x+y > 1 \right\} \\ &= \left\{ \left(\frac{1-y}{1-(x+y)}, \frac{1-x}{1-(x+y)} \right) : (x, y) \in \mathbb{X}, x+y < 1 \right\} \\ &= \{(\alpha, \beta) : \alpha \geq 1, \beta \geq 1\}, \end{aligned}$$

we see that for any $\{(\varepsilon_k, n_k, m_k), 1 \leq k \leq l\}$, $\varepsilon_k = +1$ or -1 , $n_k, m_k \geq 1$, we have

$$(2) \quad T^l\{(x, y) \in \mathbb{X} : \varepsilon_k(x, y) = \varepsilon_k, n_k(x, y) = n_k, m_k(x, y) = m_k, 1 \leq k \leq l\} = \mathbb{X} \quad \text{a.e.}$$

3. Multi-dimensional maps

In this section, we summarize results of [7], which we shall use in the following section.

We consider a map T of a bounded domain \mathbb{X} of \mathbb{R}^d onto itself with its countable partition $\mathcal{Q} = \{X_a : a \in I\}$. We assume the following:

- (i) Each X_a is a measurable and connected subset of \mathbb{X} with piecewise smooth boundary.
- (ii) There exists a finite number of subsets of \mathbb{X} , $U_0 (= \mathbb{X})$, U_1, \dots, U_N such that U_j , $1 \leq j \leq N$ are sets of positive measure and for any $a_1, \dots, a_n \in I$

$$T^n(X_{a_1} \cap T^{-1}X_{a_2} \cap \cdots \cap T^{-(n-1)}X_{a_n}) = U_j \quad \text{a.e.}$$

for some j , $0 \leq j \leq N$ whenever $X_{a_1} \cap T^{-1}X_{a_2} \cap \cdots \cap T^{-(n-1)}X_{a_n}$ is a set of positive Lebesgue measure.

(iii) For any $a \in I$, $T|_{X_a}$, the restriction of T to X_a , is injective and C^1 .

We write

$$\langle a_1, \dots, a_n \rangle = X_{a_1} \cap T^{-1}X_{a_2} \cap \cdots \cap T^{-(n-1)}X_{a_n}$$

which we call a cylinder set (of length n). We only consider cylinder sets of positive Lebesgue measure. From (iii), the restriction of T to $\langle a_1, \dots, a_n \rangle$ is injective, we can define $(T|_{\langle a_1, \dots, a_n \rangle})^{-1}$ of U_j for some j , $0 \leq j \leq N$, onto $\langle a_1, \dots, a_n \rangle$, which we denote by $\Psi_{\langle a_1, \dots, a_n \rangle}$. We fix a constant $C \geq 1$ and define the set of “Rényi cylinders” by

$$R(T) = \left\{ \langle a_1, \dots, a_n \rangle : \sup_{x \in T^n \langle a_1, \dots, a_n \rangle} |\det D\Psi_{\langle a_1, \dots, a_n \rangle}(x)| \leq C \cdot \inf_{x \in T^n \langle a_1, \dots, a_n \rangle} |\det D\Psi_{\langle a_1, \dots, a_n \rangle}(x)|, n \geq 1 \right\}.$$

Moreover we put

$$\mathcal{D}_n = \{ \langle a_1, \dots, a_n \rangle : \langle a_1, \dots, a_j \rangle \notin R(T) \text{ for } 1 \leq j \leq n \},$$

$$\mathbf{D}_n = \bigcup_{\langle a_1, \dots, a_n \rangle \in \mathcal{D}_n} \langle a_1, \dots, a_n \rangle,$$

$$\mathcal{B}_n = \{ \langle a_1, \dots, a_n \rangle \in R(T) : \langle a_1, \dots, a_{n-1} \rangle \in \mathcal{D}_{n-1} \},$$

and

$$\mathbf{B}_n = \bigcup_{\langle a_1, \dots, a_n \rangle \in \mathcal{B}_n} \langle a_1, \dots, a_n \rangle.$$

Then we consider the following conditions:

(C.1) (T, \mathcal{Q}) separates points, that is, for any $x, x' \in \mathbb{X}$ there exists $n \geq 0$ such that $T^n(x)$ and $T^n(x')$ are not the same elements in \mathcal{Q} .

(C.2) For each j , $0 \leq j \leq N$, there exists $\langle a_1, \dots, a_{s_j} \rangle \subset U_j$ such that $\langle a_1, \dots, a_{s_j} \rangle \in R(T)$ and $T^{s_j} \langle a_1, \dots, a_{s_j} \rangle = \mathbb{X}$.

(C.3) If $\langle a_1, \dots, a_n \rangle \in R(T)$, then $\langle b_1, \dots, b_m, a_1, \dots, a_n \rangle \in R(T)$ unless $\langle b_1, \dots, b_m, a_1, \dots, a_n \rangle$ is a set of Lebesgue measure 0.

(C.4)

$$\sum_{n=1}^{\infty} \lambda(\mathbf{D}_n) < \infty$$

where λ denotes d-dimensional Lebesgue measure.

$$(C.4)^* \sum_{n=1}^{\infty} \lambda(\mathbf{D}_n) \cdot \log n < \infty.$$

$$(C.5) \text{ For any } n \geq 1,$$

$$\sum_{m=0}^{\infty} \left(\sum_{\langle k_1, \dots, k_m \rangle} \left(\sup_{y \in T^m \langle k_1, \dots, k_m \rangle \cap (\bigcup_{j=1}^n \mathbf{B}_j)} |\det D\Psi_{\langle k_1, \dots, k_m \rangle}(y)| \right) \right) < +\infty.$$

$$(C.6) \# \mathcal{D}_1 < \infty.$$

$$(C.7) \text{ There exists a positive integer } l \text{ such that for all } n > 0 \text{ and all } \langle a_1, \dots, a_n \rangle \in \mathcal{D}_n,$$

$$\frac{\sup_{x \in T^n \langle a_1, \dots, a_n \rangle} |\det D\Psi_{\langle a_1, \dots, a_n \rangle}(x)|}{\inf_{x \in T^n \langle a_1, \dots, a_n \rangle} |\det D\Psi_{\langle a_1, \dots, a_n \rangle}(x)|} = O(n^l).$$

$$(C.8) \log |\det DT(\cdot)| \text{ is Lebesgue integrable.}$$

$$(C.9) \text{ there exists a positive integer } k_0 \text{ such that if } \langle a_1, \dots, a_n \rangle \in \mathcal{D}_n^c \text{ and } \langle a_2, \dots, a_n \rangle \in \mathcal{D}_{n-1}, \text{ then}$$

$$\langle a_1, \dots, a_n \rangle \subset \bigcup_{j=1}^{k_0} \mathbf{B}_j.$$

Then we have the following.

Theorem 3.1 ([7]). (i) (C.1)–(C.4) imply that there exists an absolutely continuous invariant probability measure μ and (T, μ) is exact, i.e.

$$\bigcap_{k=1}^{\infty} T^{-k} \mathbb{B}$$

is trivial, where \mathbb{B} denotes the set of Borel subsets of \mathbb{X} .

(ii) (C.1)–(C.8) imply Rohlin's entropy formula:

$$h(T) = \int_{\mathbb{X}} \log |\det DT(x)| d\mu(x).$$

(iii) (C.1)–(C.9) with (C.4)* imply that (T, μ, \mathcal{Q}) is weak Bernoulli, that is, for any $\varepsilon > 0$, there exists $n_0 > 0$ such that $\{\Delta_k\}$ and $\{\Delta_l\}$ are ε -independent for any $k \geq 1$, $l \geq k + n$, and $n \geq n_0$, where $\{\Delta_k\}$ and $\{\Delta_l\}$ denote the sets of cylinder sets of length k and l respectively and Two partitions \mathcal{F}_1 and \mathcal{F}_2 are said to be ε -independent if

$$\sum_{A \in \mathcal{F}_1} \sum_{B \in \mathcal{F}_2} |\mu(A \cap B) - \mu(A)\mu(B)| < \varepsilon.$$

4. Ergodic properties of the negative slope algorithm

First of all, from (2), we can take $\{U_0\}$ as $\{U_0, \dots, U_N\}$ in the previous section ($U_0 = \mathbb{X}$). We show the following.

Theorem 4.1. *There exists an absolutely continuous invariant probability measure μ for T and (T, μ) is exact.*

REMARK 4.2. [4] discussed the explicit form of the density function $d\mu/d\lambda$, which we will see later, and showed its ergodicity. The exactness implies not only ergodicity but also mixing of all degrees.

To prove this theorem, we will show that T satisfies the conditions (C.1)–(C.4). We define the set $R(T)$ by

$$R(T) = \{(\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k) \mid (\varepsilon_k, n_k, m_k) \neq (+1, 1, 1)\}.$$

In the sequel, we simply write Δ_k for a cylinder set

$$((\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k))$$

if it is clear in the context. We put

$$A_{(\varepsilon_1, n_1, m_1)}^{-1} \cdots A_{(\varepsilon_k, n_k, m_k)}^{-1} = \begin{pmatrix} p_1^{(k)} & p_2^{(k)} & p_3^{(k)} \\ r_1^{(k)} & r_2^{(k)} & r_3^{(k)} \\ q_1^{(k)} & q_2^{(k)} & q_3^{(k)} \end{pmatrix}$$

for any sequence $((\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k))$, $k \geq 1$. Then it is easy to see that $q_1^{(k)} = q_2^{(k)}$.

Lemma 4.3. *For any sequence $((\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k))$, $\varepsilon_i = \pm 1$, $n_i, m_i \geq 1$, $1 \leq i \leq k$, we see*

- (i) $T^k(\Delta_k) = \mathbb{X}$,
- (ii)

$$|\det D\Psi_{\Delta_k}(x, y)| = \frac{1}{(q_1^{(k)}x + q_2^{(k)}y + q_3^{(k)})^3}.$$

Proof. It is an easy consequence of induction and calculation, respectively, see also F. Schweiger [6], Proposition 2 for (ii). \square

From this lemma, it is easy to see the following.

Lemma 4.4. *If $\langle(\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k)\rangle \in R(T)$, then*

$$\sup_{(x,y) \in \mathbb{X}} |\det D\Psi_{\Delta_k}(x, y)| \leq 3^3 \inf_{(x,y) \in \mathbb{X}} |\det D\Psi_{\Delta_k}(x, y)|.$$

Therefore, $R(T)$ is the set of Rényi cylinders.

Proof. Since

$$A_{(\varepsilon_1, n_1, m_1)}^{-1} \cdots A_{(\varepsilon_k, n_k, m_k)}^{-1} = \begin{cases} \begin{pmatrix} p_1^{(k-1)} & p_2^{(k-1)} & p_3^{(k-1)} \\ r_1^{(k-1)} & r_2^{(k-1)} & r_3^{(k-1)} \\ q_1^{(k-1)} & q_2^{(k-1)} & q_3^{(k-1)} \end{pmatrix} \begin{pmatrix} 1 & 0 & n_k - 1 \\ 0 & 1 & m_k - 1 \\ 1 & 1 & n_k + m_k - 1 \end{pmatrix} & \text{if } \varepsilon_k = +1 \\ \begin{pmatrix} p_1^{(k-1)} & p_2^{(k-1)} & p_3^{(k-1)} \\ r_1^{(k-1)} & r_2^{(k-1)} & r_3^{(k-1)} \\ q_1^{(k-1)} & q_2^{(k-1)} & q_3^{(k-1)} \end{pmatrix} \begin{pmatrix} 0 & 1 & m_k \\ 1 & 0 & n_k \\ 1 & 1 & n_k + m_k - 1 \end{pmatrix} & \text{if } \varepsilon_k = -1 \end{cases},$$

we see that

$$(q_1^{(k)}, q_2^{(k)}, q_3^{(k)}) = \begin{cases} (q_1^{(k-1)} + q_3^{(k-1)}, q_2^{(k-1)} + q_3^{(k-1)}, (n_k - 1)q_1^{(k-1)} + (m_k - 1)q_2^{(k-1)} + (n_k + m_k - 1)q_3^{(k-1)}) & \text{if } \varepsilon_k = +1 \\ (q_2^{(k-1)} + q_3^{(k-1)}, q_1^{(k-1)} + q_3^{(k-1)}, m_k q_1^{(k-1)} + n_k q_2^{(k-1)} + (n_k + m_k - 1)q_3^{(k-1)}) & \text{if } \varepsilon_k = -1. \end{cases}$$

It follows by induction that $q_1^{(k)} = q_2^{(k)}$ for any $k \geq 1$. If $\Delta_k \in R(T)$, then $n_k + m_k \geq 3$ or $n_k + m_k \geq 2$ when $\varepsilon_k = -1$ or $+1$, respectively. Thus we see $q_i^{(k)} < q_3^{(k)}$, $i = 1, 2$, whenever $\Delta_k \in R(T)$. By Lemma 4.3, we have

$$\frac{1}{(q_1^{(k)} + q_2^{(k)} + q_3^{(k)})^3} \leq |\det D\Psi_{\Delta_k}(x, y)| \leq \frac{1}{(q_3^{(k)})^3}.$$

Hence we get

$$\frac{1}{(3q_3^{(k)})^3} \leq |\det D\Psi_{\Delta_k}(x, y)| \leq \frac{1}{(q_3^{(k)})^3},$$

which implies the assertion of this lemma. \square

Let's define the following:

$$\mathcal{D}_k = \{ \langle(\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k)\rangle \mid \langle(\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_i, n_i, m_i)\rangle \notin R(T) \text{ for } 1 \leq i \leq k \},$$

$$\mathbf{D}_k = \bigcup_{\Delta_k \in \mathcal{D}_k} \Delta_k,$$

$$\mathcal{B}_k = \{ \langle (\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k) \rangle \in R(T) \mid \\ \langle (\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_{k-1}, n_{k-1}, m_{k-1}) \rangle \in \mathcal{D}_{k-1} \},$$

and

$$\mathbf{B}_k = \bigcup_{\Delta_k \in \mathcal{B}_k} \Delta_k.$$

It is easy to see that

$$\mathcal{D}_k = \{ \underbrace{\langle (+1, 1, 1), \dots, (+1, 1, 1) \rangle}_{k \text{ times}} \}.$$

Now we will check the conditions of [7]. First of all, it is clear that the set of cylinder sets separates points, see (1). Lemma 4.3 (i) and Lemma 4.4 imply (C.2) and (C.3), respectively. We see the following.

Lemma 4.5 ((C.4)). *We have*

$$\sum_{k=1}^{\infty} \lambda(\mathbf{D}_k) < \infty$$

where λ denotes the 2-dimensional Lebesgue measure.

Proof. From the definition of T and simple calculation, we see that

$$\langle (+1, 1, 1) \rangle = \left\{ (x, y) \mid 0 \leq y < 1 - 2x, 0 \leq y < \frac{1}{2} - \frac{1}{2}x \right\}$$

and, in general,

$$\underbrace{\langle (+1, 1, 1), \dots, (+1, 1, 1) \rangle}_{k \text{ times}} \\ = \left\{ (x, y) \mid 0 \leq y < \frac{1}{k} - \frac{k+1}{k}x, 0 \leq y < \frac{1}{k+1} - \frac{k}{k+1}x \right\}.$$

Hence we have

$$\lambda(\mathbf{D}_k) = \frac{1}{(k+1)(2k+1)}$$

and get the conclusion of this lemma. □

This completes the proof of Theorem 4.1 by [7].

Next we show the following.

Theorem 4.6 (Rohlin's formula). *The entropy $H_\mu(T)$ of (\mathbb{X}, T, μ) is given by*

$$H_\mu(T) = \int_{\mathbb{X}} \log |\det DT| d\mu.$$

In the following, we show (C.5)–(C.8) in [7], which imply this theorem.

Lemma 4.7 ((C.5)).

$$W_k = \sum_{l=0}^{\infty} \sum_{\Delta_l \in \mathcal{D}_l} \left(\sup_{(x,y) \in (\bigcup_{j=1}^k \mathbf{B}_j)} |\det D\Psi_{\Delta_l}(x, y)| \right) < \infty.$$

Proof. Since $\Delta_l \in \mathcal{D}_l$ means

$$\Delta_l = \underbrace{(+1, 1, 1), \dots, (+1, 1, 1)}_{l \text{ times}},$$

Ψ_{Δ_l} is associated to

$$\Psi_{\Delta_l} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}}_{l \text{ times}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l & l & 1 \end{pmatrix}.$$

Thus we see

$$(3) \quad \det D\Psi_{\Delta_l}(x, y) = \frac{1}{(lx + ly + 1)^3}.$$

On the other hand,

$$\bigcup_{j=1}^k \mathbf{B}_j = \mathbb{X} \setminus \Delta_k$$

and

$$\min_{(x,y) \in \bigcup_{j=1}^k \mathbf{B}_j} x + y = \frac{1}{k+1},$$

see the proof of Lemma 4.5. Hence we get

$$\begin{aligned} \sup_{(x,y) \in (\bigcup_{j=1}^k \mathbf{B}_j)} |\det D\Psi_{\Delta_l}(x, y)| &= \frac{1}{((1/(k+1))l+1)^3} \\ &\leq (k+1)^3 \frac{1}{l^3}, \end{aligned}$$

which implies the assertion of this lemma. \square

Lemma 4.8 ((C.6)).

$$\sharp \mathcal{D}_1 = 1.$$

Proof. This is obvious. \square

Lemma 4.9 ((C.7)). *We have*

$$\frac{\sup_{(x,y) \in \mathbb{X}} |\det D\Psi_{\Delta_k}(x, y)|}{\inf_{(x,y) \in \mathbb{X}} |\det D\Psi_{\Delta_k}(x, y)|} = \mathcal{O}(k^3)$$

for $\Delta_k = \underbrace{\langle (+1, 1, 1), \dots, (+1, 1, 1) \rangle}_{k \text{ times}}.$

Proof. This follows from (3). \square

Lemma 4.10 ((C.8)). *The function $\log|\det DT|$ is integrable with respect to λ .*

Proof.

$$\begin{aligned} &\int_{\mathbb{X}} \log|\det DT| d\lambda \\ &= - \iint_{\mathbb{X} \cap \{x+y>1\}} 3 \log((x+y)-1) dx dy - \iint_{\mathbb{X} \cap \{x+y<1\}} 3 \log(1-(x+y)) dx dy. \end{aligned}$$

Then, there exists $K > 0$ s.t.

$$\int_{\mathbb{X}} \log|\det DT| d\lambda < K \int_0^2 \log r dr < \infty.$$

This completes the proof of the Theorem 4.6. \square

Finally, we show the following.

Theorem 4.11. *The negative slope algorithm with the absolutely continuous invariant probability measure μ is weak Bernoulli.*

To prove this theorem we need the following two lemmas.

Lemma 4.12 ((C.4)*).

$$\sum_{k=1}^{\infty} \lambda(\mathbf{D}_k) \cdot \log k < \infty.$$

Proof. This is clear since $\lambda(\mathbf{D}_k) = 1/((k+1)(2k+1))$, see the proof of Lemma 4.5. \square

Lemma 4.13 ((C.9)). *If $\langle(\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k)\rangle \in \mathcal{D}_k^c$ and $\langle(\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k)\rangle \in \mathcal{D}_{k-1}$, then we have $\langle(\varepsilon_1, n_1, m_1)\rangle \in \mathbf{B}_1$, that is, $(\varepsilon_1, n_1, m_1) \neq (+1, 1, 1)$.*

Proof. This is an easy consequence of the definitions of \mathcal{D}_k and \mathcal{B}_k . \square

Since T satisfies (C.1)–(C.9) with (C.4)*, we can conclude the assertion of Theorem 4.11.

5. Absolutely continuous invariant measure

In [4], it was shown that the density function of the absolutely continuous invariant probability measure was given by

$$\frac{d\mu}{d\lambda} = \frac{1}{2 \log 2} \frac{1}{x+y}.$$

This was checked by Kuzmin's equation

$$f(x, y) = \sum_{\varepsilon=\pm 1, n, m \geq 1} f(\Psi_{(\varepsilon, n, m)}(x, y)) |\det \Psi_{(\varepsilon, n, m)}(x, y)|$$

for $f(x, y) = 1/(x+y)$.

In the sequel, we prove the same result by a different way, which is called a “natural extension method” originally started by [5] for a class of continued fraction transformations. We start with a 4-dimensional area. Let $\overline{\mathbb{X}} = \mathbb{X} \times (-\infty, 0)^2$. For $(x, y, z, w) \in \overline{\mathbb{X}}$, we define a map \overline{T} on $\overline{\mathbb{X}}$ by

$$\overline{T}(x, y, z, w) = \begin{cases} \left(\frac{y}{(x+y)-1} - n(x, y), \frac{x}{(x+y)-1} - m(x, y), \frac{w}{(z+w)-1} - n(x, y), \frac{z}{(z+w)-1} - m(x, y) \right) & \text{if } x+y > 1 \\ \left(\frac{1-y}{1-(x+y)} - n(x, y), \frac{1-x}{1-(x+y)} - m(x, y), \frac{1-w}{1-(z+w)} - n(x, y), \frac{1-z}{1-(z+w)} - m(x, y) \right) & \text{if } x+y < 1. \end{cases}$$

Then it is easy to see that \overline{T} is bijective on $\overline{\mathbb{X}}$ except for the set of (4-dimensional) Lebesgue measure 0.

Proposition 5.1. *The measure $\overline{\mu}$ defined by*

$$\frac{d\overline{\mu}}{d\overline{\lambda}} = \frac{1}{\{(x+y) - (z+w)\}^3}$$

is an invariant measure for \overline{T} , where $\overline{\lambda}$ denotes the 4-dimensional Lebesgue measure.

Proof. We put

$$h(x, y, z, w) = \frac{1}{\{(x+y) - (z+w)\}^3}.$$

It is enough to show that

$$h(\overline{T}(x, y, z, w)) \cdot |\det D(\overline{T}(x, y, z, w))| \cdot h^{-1}(x, y, z, w) = 1,$$

which follows easily by simple calculation.

CASE i. $x + y > 1$

$$\begin{aligned} & h(\overline{T}(x, y, z, w)) \cdot |\det D(\overline{T}(x, y, z, w))| \cdot h^{-1}(x, y, z, w) \\ &= \frac{1}{((y+x)/((x+y)-1) - (w+z)/((z+w)-1))^3} \\ (4) \quad & \cdot \left| \frac{1}{(x+y-1)^3(z+w-1)^3} \right| \cdot \left(\frac{1}{((x+y) - (z+w))^3} \right)^{-1} \\ &= 1. \end{aligned}$$

CASE ii. $x + y < 1$

$$\begin{aligned} & h(\overline{T}(x, y, z, w)) \cdot |\det D(\overline{T}(x, y, z, w))| \cdot h^{-1}(x, y, z, w) \\ &= \frac{1}{(((1-y) + (1-x))/(1-(x+y)) - ((1-w) + (1-z))/(1-(z+w)))^3} \\ (5) \quad & \cdot \left| \frac{1}{(1-(x+y))^3(1-(z+w))^3} \right| \cdot \left(\frac{1}{(((1-(x+y)) - (1-(z+w)))^3} \right)^{-1} \\ &= 1. \end{aligned}$$

(4) and (5) imply the assertion of this proposition. □

Corollary 5.2. *The measure μ defined by*

$$\frac{d\mu}{d\lambda} = \frac{1}{2 \log 2} \frac{1}{(x+y)}$$

is an invariant probability measure for T .

Proof. It is easy to see that the projection of $\overline{\mu}$ to \mathbb{X} is an invariant measure for T . We have

$$\int_{(-\infty, 0) \times (-\infty, 0)} \frac{1}{\{(x+y) - (z+w)\}^3} dz dw = \frac{1}{2} \frac{1}{(x+y)},$$

which is the assertion of this corollary. \square

From this formula, we can compute the entropy $H_\mu(T)$ from Theorem 4.6.

Proposition 5.3.

$$H_\mu(T) = \frac{\pi^2}{4 \log 2}.$$

Proof. From Theorem 4.6, we have

$$\begin{aligned} H_\mu(T) &= -\frac{3}{2 \log 2} \iint_{\{x+y>1\}} \frac{1}{x+y} \log((x+y) - 1) dx dy \\ &\quad - \frac{3}{2 \log 2} \iint_{\{x+y<1\}} \frac{1}{x+y} \log(1 - (x+y)) dx dy. \end{aligned}$$

The right side is equal to

$$\begin{aligned} &-\frac{3}{2 \log 2} \left[\int_1^2 \frac{2-t}{t} \log(t-1) dt + \int_0^1 \log(1-t) dt \right] \\ &= -\frac{3}{\log 2} \left[\int_0^1 \frac{1}{1+t} \log t dt \right] \\ &= \frac{\pi^2}{4 \log 2}. \end{aligned} \quad \square$$

From this proposition, we can get the exponential divergence of $q_3^{(k)}$ as $k \rightarrow \infty$.

Proposition 5.4.

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log q_3^{(k)} = \frac{\pi^2}{12 \log 2}$$

for λ -a.e. (x, y) .

Proof. From the Shannon-MacMillan-Breiman theorem, we have

$$-\lim_{k \rightarrow \infty} \frac{1}{k} \log \mu(\Delta_k) = \frac{\pi^2}{4 \log 2} \quad \mu\text{-a.e.}$$

where Δ_k is defined by $(\varepsilon_i, n_i, m_i) = (\varepsilon_i(x, y), n_i(x, y), m_i(x, y))$ for $1 \leq i \leq k$. We take (x, y) so that (4) holds. Then we choose a subsequence $(l_k): k \geq 1$ by

$$l_1 = \min\{l \geq 1 \mid (\varepsilon_l(x, y), n_l(x, y), m_l(x, y)) \neq (+1, 1, 1)\}$$

and

$$l_{k+1} = \min\{l > l_k \mid (\varepsilon_l(x, y), n_l(x, y), m_l(x, y)) \neq (+1, 1, 1)\}$$

for $k \geq 1$, which means that we choose all cylinders $\Delta_l \in R(T)$. Since Δ_l is bounded away from 0, there exists a constant $C_1 > 1$ such that

$$\frac{1}{C_1} \lambda(\Delta_{l_k}) < \mu(\Delta_{l_k}) < C_1 \lambda(\Delta_{l_k}).$$

On the other hand, there exists a constant $C_2 > 1$ such that

$$\frac{1}{C_2 q_3^{(l)}} < \lambda(\Delta_l) < \frac{C_2}{q_3^{(l)}}$$

whenever $\Delta_l \in R(T)$, see Lemma 4.3. Thus we get

$$\lim_{k \rightarrow \infty} \frac{1}{l_k} \log q_3^{(l_k)} = \frac{\pi^2}{12 \log 2}$$

for μ -a.e. (x, y) . It is clear that $q_3^{(k+1)} = q_3^{(k)}$ if $(\varepsilon_k(x, y), n_k(x, y), m_k(x, y)) = (+1, 1, 1)$. Since the indicator function of $\langle (+1, 1, 1) \rangle$ is obviously integrable with respect to μ ,

$$\lim_{k \rightarrow \infty} \frac{l_k - l_{k-1}}{l_k} = 0$$

for μ -a.e. (x, y) . Hence we have

$$\lim_{l \rightarrow \infty} \frac{1}{l} \log q_3^{(l)} = \frac{\pi^2}{12 \log 2}$$

for μ -a.e. (x, y) , equivalently λ -a.e. \square

REMARK 5.5. It is easy to see that

$$\left(\frac{p_3^{(k)}}{q_3^{(k)}}, \frac{r_3^{(k)}}{q_3^{(k)}} \right), \quad \left(\frac{p_1^{(k)} + p_3^{(k)}}{q_1^{(k)} + q_3^{(k)}}, \frac{r_1^{(k)} + r_3^{(k)}}{q_1^{(k)} + q_3^{(k)}} \right), \quad \left(\frac{p_2^{(k)} + p_3^{(k)}}{q_2^{(k)} + q_3^{(k)}}, \frac{r_2^{(k)} + r_3^{(k)}}{q_2^{(k)} + q_3^{(k)}} \right),$$

and

$$\left(\frac{p_1^{(k)} + p_2^{(k)} + p_3^{(k)}}{q_1^{(k)} + q_2^{(k)} + q_3^{(k)}}, \frac{r_1^{(k)} + r_2^{(k)} + r_3^{(k)}}{q_1^{(k)} + q_2^{(k)} + q_3^{(k)}} \right)$$

are $\Psi_{\Delta_k}(0, 0)$, $\Psi_{\Delta_k}(1, 0)$, $\Psi_{\Delta_k}(0, 1)$, and $\Psi_{\Delta_k}(1, 1)$, respectively. Then it is also possible to show that

$$\lim_{k \rightarrow \infty} \frac{1}{k} (\log q_1^{(k)} + \log q_2^{(k)} + \log q_3^{(k)}) = \frac{\pi^2}{12 \log 2}$$

for λ -a.e. (x, y) by the same way. We note that these four sequences converge to (x, y) because of (C.1). Suppose that

$$d(k, x, y) = \text{diameter}(\langle (\varepsilon_1, n_1, m_1), (\varepsilon_2, n_2, m_2), \dots, (\varepsilon_k, n_k, m_k) \rangle).$$

Then the convergence rate of the above four sequences to (x, y) is bounded by $d(k, x, y)$.

ACKNOWLEDGEMENT. The authors would like to thank the referee of this manuscript for his careful reading.

References

- [1] S. Ferenczi, C. Holton and L.Q. Zamboni: *Structure of three interval exchange transformations I, an arithmetic study*, Ann. Inst. Fourier (Grenoble) **51** (2001), 861–901.
- [2] S. Ferenczi, C. Holton and L.Q. Zamboni: *Structure of three-interval exchange transformations II, a combinatorial description of the trajectories*, J. Anal. Math. **89** (2003), 239–276.
- [3] S. Ferenczi, C. Holton and L. Q. Zamboni: *Structure of three-interval exchange transformations III, ergodic and spectral properties*, J. Anal. Math. **93** (2004), 103–138.
- [4] S. Ferenczi and L.F.C. da Rocha: *A self-dual induction for three-interval exchange transformations*, preprint.
- [5] H. Nakada: *Metrical theory for a class of continued fraction transformations and their natural extensions*, Tokyo J. Math. **4** (1981), 399–426.
- [6] F. Schweiger: *Multidimensional Continued Fractions*, Oxford Univ. Press, Oxford, 2000.
- [7] M. Yuri: *On a Bernoulli property for multidimensional mappings with finite range structure*, Tokyo J. Math. **9** (1986), 457–485.

Koshiro Ishimura
Department of Mathematics
Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama
Japan
e-mail: koshiro@math.keio.ac.jp

Hitoshi Nakada
Department of Mathematics
Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama
Japan
e-mail: nakada@math.keio.ac.jp